Expert Systems With Applications 122 (2019) 369–375

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Minimizing greenhouse gas emissions using inverse DEA with an application in oil and gas

Matthew Wegener, Gholam R. Amin*

Faculty of Business, University of New Brunswick at Saint John, Saint John, New Brunswick E2L 4L5, Canada

ARTICLE INFO

Article history:
Received 12 August 2018
Revised 6 November 2018
Accepted 31 December 2018
Available online 31 December 2018

Keywords:
Data envelopment analysis
Inverse DEA
Undesirable outputs
Greenhouse gas emissions
Oil and gas

ABSTRACT

The industry dynamic creates a difficult problem that must be solved if we are to address climate change. How can we lower greenhouse gas (GHG) emissions while simultaneously increasing production? This paper develops a new inverse data envelopment analysis (DEA) model for optimizing GHG emissions. The inverse DEA model minimizes the overall GHG emissions generated by a set of decision making units (DMUs) for producing a certain level of outputs, given that the DMUs maintain at least their existing performance status. The usefulness of the proposed method in this paper is demonstrated through an application in the oil and gas industry. We find that roughly 57% of our sample DMUs are inefficient. Ample room exists within the current efficiency frontier to lower GHG emissions in the oil and gas sector. We recommend that environmental regulators look into targeting regulations at these inefficiencies that exist within the industry as low bearing fruit for potential GHG emissions reductions.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

On November 4, 2016 the Paris Agreement on climate change came into force. While the agreement requires reductions in greenhouse gases (GHG) at the national level, it is safe to say that at least part of the burden will be carried by firms in high GHG producing industries. We are effectively at the beginning of a carbon constrained global economy. Yet some of these products, like oil and gas, are so integral to our societies, that it is difficult to see how we can get by without them. Rather than suggest changes in consumption, we consider the impact of firms including GHG emissions as an additional production output when comparing efficiency. Ultimately, we seek to determine how additional production should be distributed between firms such that the industry's overall GHG emissions will be minimized. Despite the inclusion of GHG emissions as an additional production output, we maintain that firms should be able to at least maintain their current performance status. This paper develops an inverse data envelopment analysis (DEA) model to solve this problem.

The energy sector is considered to be the largest contributor to global GHG emissions. According to the Intergovernmental Panel on Climate Change, the energy sector, which includes oil and gas, was responsible for roughly 35% of the global anthropogenic GHG emissions in 2010.[1] Not only does the sector account for a large percentage of GHG emissions, but the rate of growth in GHG emissions has been accelerating. The growth rate in GHG emissions for the sector rose from 1.7% per year between 1990 and 2000 to 3.1% per year from 2000 to 2010[1]. The sector's contribution to climate change has been increasing when the Canadian government is targeting a 30% reduction in GHG emissions, relative to its 2005 GHG emissions levels, by 2030. This targeted reduction comes at a time when Canadian oil production is expected to grow by 35% between 2017 and 2035.[2] This industry dynamic creates a difficult problem that must be solved if we are to address climate change. How can we lower GHG emissions while simultaneously increasing production? Our application of the inverse DEA model we propose is a first step to finding a win-win solution. We seek to find a way to address our environmental needs without harming our economy.

We contribute to the literature by developing a novel real world application of an inverse DEA model. The proposed inverse DEA model in this paper can estimate the minimum GHG emissions for an increase in production for a chosen set of firms with the assumption that the firms' efficiency will not be deteriorated. We apply our proposed model to Canadian and US integrated oil and gas companies. In doing so, we reveal that there is ample room within

[1] IPCC (2014). Climate Change 2014 Mitigation of Climate Change. <http://www.ipcc.ch/report/ar5/wg3/>. Retrieved: 8/7/2008.

[2] CBC (2018). Canadian total oil production expected to rise by a third by 2035. <https://www.cbc.ca/news/business/oil-production-canada-1.4702433>. Retrieved: 8/7/2018.

* Corresponding author.
E-mail addresses: mwegener@unb.ca (M. Wegener), gamin@unb.ca (G.R. Amin).

https://doi.org/10.1016/j.eswa.2018.12.058
0957-4174/© 2018 Elsevier Ltd. All rights reserved.

the current efficiency frontier to increase oil production without increasing GHG emissions. This should be of great importance to both managers of oil and gas companies and environmental policy setters. For managers of inefficient firms, a unique opportunity exists. If they are willing to improve their eco-efficiency, they have room to produce more barrels of oil equivalent without increasing their GHG emissions. They have room to grow in a carbon constrained economy, but they also risk high regulatory costs if they do not change their behaviour. For environmental regulators, our inverse DEA model can provide indication of how much each company can increase their production without impacting GHG emissions significantly. This represents low bearing fruit for GHG emissions improvement. Regulators should seek to incentivise improving eco-efficiency in the inefficient subset of firms. We recommend regulators consider fining, or taxing companies based on their potential to increase production free of GHG emissions increases. This would be a regulatory avenue that could help address climate change without forcing consumers to change their behaviour through additional carbon taxes.

The remainder of the paper is organized as follows. First we examine the literature related to the use of DEA in environmental efficiency as well as the literature of the inverse DEA. The literature review is followed by the development of an inverse DEA model to determine optimal GHG emissions in Section 3. Section 4 explains the application consisting of oil and gas firms, data collection and selection of inputs and outputs. The developed model is then applied to Canadian and US integrated oil and gas companies in Section 5. We discuss the results in Section 6 and the final section is concluding remarks with suggestions for future research.

2. Literature review

We examine the literature review in two parts. In the first part, we review the literature of DEA in environmental efficiency. We also review the inverse DEA and its most recent development theory and application.

2.1. DEA in environmental efficiency

In this section, we discuss the prior literature that has used DEA to assess environmental efficiency. As a non-parametric performance evaluation method, DEA is well known a successful method which has been widely used in environmental efficiency.

Sueyoshi and Wang (2014) used DEA to assess the operational and environmental efficiencies of petroleum firms in the United States. They examined the differences in DEA assessment results between large integrated oil and gas firms and small independent firms that were restricted to upstream operations. Sueyoshi and Wang (2014) concluded that large integrated oil and gas firms were able to outperform their small competitors. Sueyoshi and Goto (2014a) used DEA to compare photovoltaic power stations efficiency between Germany and the United States. Since photovoltaic power stations do not produce GHG emissions, they represent a means to increase environmental efficiency in the energy sector. They concluded that Germany outperformed the United States despite solar and land disadvantages. They attributed this to effective energy policy in Germany, but warned that the improvements came at a cost to consumers. Sueyoshi and Goto (2014b) incorporated the concepts of natural and managerial disposability into a DEA model. These two strategic concepts represent the desire to lower GHG emissions and to increase financial performance respectively. Thus, Sueyoshi and Goto (2014b) introduced the use of DEA to optimize not only desirable but also undesirable performance. They used a radial model that measured unified efficiency. They illustrated this use of DEA by comparing Japanese chemical and pharmaceutical firms. Zhou, Chung, and Zhang (2014) examined the energy efficiency of Chinese transport sector. They suggested a DEA approach with the ability to include undesirable outputs to maximize the energy-saving potential of the transport sector. Mukherjee (2010) applied the directional distance function to measure energy efficiency in the manufacturing sector in India. Mukherjee (2010) concluded that states should be able to increase output while reducing energy inputs by an annual average of 3.84% through technical efficiency improvements. Molinos-Senante, Hernandez-Sancho, Mocholi-Arce, and Sala-Garrido (2014) used DEA to quantify the potential for GHG reduction in Spanish wastewater treatment plants. Then, using the results from their DEA model, they isolate factors affecting wastewater treatment plants' environmental performance. Vlontzos, Niavis, and Manos (2014) employed a non-radial DEA model to evaluate the energy and environmental efficiency of the primary sectors of the EU member state countries. Riccardi, Oggioni, and Tonineli (2012) used both DEA models and a directional distance function approach to determine cement sector efficiency. They assessed the efficiency both with consideration of GHG emissions and without. They found that GHG emissions, investments in new technologies, and the utilization of alternative fuels and raw materials impacted the efficiency of the cement production processes.

In all of the above studies, the DEA method has been used for assessing the environmental efficiency. The classic DEA models do not reveal the levels of inputs and outputs that would allow firms to realize given efficiency targets. Therefore, solving inverse problems is beyond the scope of the DEA method and requires appropriate inverse models. Being completely two different concepts, the results are not really comparable, inverse DEA focuses on the efficiency and finds the data point while DEA focuses on the data and finds the efficiency score.

In this paper, we introduce a novel inverse problem in environmental efficiency and develop an inverse DEA model to solve it.

2.2. Inverse DEA

This section briefly reviews the origin, development and recent advances in inverse DEA. Wei, Zhang, and Zhang (2000) developed the first inverse DEA model to solve some novel inverse problems in DEA. Then, Yan, Wei, and Hao (2002) further developed the inverse DEA by applying it in the resource allocation. Unlike the conventional DEA, inverse DEA solves inverse problems in the DEA context. In an inverse problem, the efficiency score of a DMU is given and the objective is to determine the optimal quantity of inputs and/or outputs that is required to reach the efficiency target. Gattoufi, Amin, and Emrouznejad (2014) introduced a novel application of inverse DEA in mergers and acquisitions to estimate the optimal level of inputs and outputs from pre-merger DMUs which is needed to be kept by the post-merger entity to reach a given efficiency target. Amin, Emrouznejad, and Gattoufi (2017) generalized the concept of mergers and acquisitions using inverse DEA. They introduced a generalized restructuring as a move in a business market in which a set of pre-restructuring firms can be converted into a set of post-restructuring firms aiming to improve the efficiency of post-restructuring entities in the forms of a consolidation using synergy or a split using reverse synergy. Recently, Amin and Al-Muharrami (2018) developed an inverse DEA model that can address inverse problems in the mergers and acquisitions of firms with negative data. Also, Emrouznejad, Yang, and Amin (2018) developed an inverse DEA model to determine the optimal way to allocate CO_2 emissions quota in the Chinese manufacturing industries.

Lim (2016) suggested an inverse DEA application for a new product target setting given expected changes of production frontier. Jahanshahloo, Soleimani-Damaneh, and Ghobadi (2015)

developed a time based inverse DEA model assuming intertemporal dependence of dataset. Lin (2010) addressed the efficiency measurement and revenue setting problems of chain stores in Taiwan by using the imprecise DEA and inverse DEA models. Lertworasirikul, Charnsethikul, and Fang (2011) presented an inverse BCC model to allocate resources with no changes in efficiency scores. Ghiyasi (2017) developed inverse DEA models based on cost and revenue efficiency when price information is available. Eyni, Tohidi, and Mehrabeian (2017) proposed an application of the inverse DEA to sensitivity analysis of DMUs with undesirable inputs and outputs. A literature review with latest development on DEA and inverse DEA can be found in Emrouznejad and Yang (2018).

The inverse DEA has the potential to solve challenging decision making problems however, there are few applications of inverse DEA in the real world. To fill this gap in the literature, this paper introduces a novel inverse DEA model to minimize GHG emissions with an application in oil and gas industry.

3. Optimizing GHG emissions

Suppose that there are n DMUs where the jth DMU consumes m inputs x_{ij} ($i=1, \ldots, m$) to produce s good or desirable outputs y^g_{rj} ($r=1, \ldots, s$) and q bad or undesirable outputs y^b_{pj} ($p=1, \ldots, q$) for each $j=1, \ldots, n$.

We use the following directional distance DEA model to obtain the efficiency score of DMU-k (for each $k=1, \ldots, n$).

$$
\begin{aligned}
&\beta^*_k = \max\ \beta \\
&\quad s.t. \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \le x_{ik} \qquad i=1,\ldots,m \\
&\sum_{j=1}^{n} \lambda_j y^g_{rj} \ge (1+\beta) y^g_{rk} \quad r=1,\ldots,s \\
&\sum_{j=1}^{n} \lambda_j y^b_{pj} = (1-\beta) y^b_{pk} \quad p=1,\ldots,q \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \ge 0 \quad j=1,\ldots,n
\end{aligned}
\tag{1}
$$

The efficiency score of DMU-k is $(1-\beta^*_k)/(1+\beta^*_k)$ Chung, Färe, & Grosskopf, 1997). Therefore, DMU-k is efficient if and only if $\beta^*_k = 0$. In fact, the optimal value of model ((1), or β^*_k, provides a measure of the inefficiency of DMU-k.

The directional distance model (1) has the advantage of imposing undesirable outputs reduction and desirable outputs expansion simultaneously. The directional distance models have been extensively used in the DEA literature. Shetty, Pakkala, and Mallikarjunappa (2012) proposed a modified directional distance DEA formulation to assess bankruptcy with an application to IT firms in India. Toloo, Allahyar, and Hančlová (2018) developed a non-radial directional distance model to classify inputs and outputs with an application in banking sector.

In this paper, we consider the directional distance DEA model (1) as the base optimization model to solve a novel inverse problem. The directional distance model is more relevant than other DEA models to our application in oil and gas sector. Many oil and gas firms have the capacity to reduce their GHG emissions for their level of production, the directional distance DEA model allows us to use this capacity while simultaneously increasing production. However, the proposed inverse DEA model in this paper can be developed for any other DEA model that can handle undesirable outputs.

The objective of this paper is to solve the following inverse problem: For a set of selected firms and a certain level of production, what is the minimum level of overall undesirable outputs that will be generated by the set of firms to produce the given quantity of desirable outputs, assuming the firms should maintain their post-production efficiency score at least the same as their pre-production efficiency score?

Let S be a set of DMUs that are selected to produce an additional production quantity, assuming their efficiency scores remain at least the same as their current efficiency scores. Moreover, let $\hat{y}^g_r$ be a given additional production quantity of the rth good output that is going to be produced by the set of DMUs in S, for each $r=1, \ldots, s$. As denoted in model (1), the inputs and outputs of DMU-k (for each $k \in S$) are $(\mathbf{x}_k, \mathbf{y}^g_k, \mathbf{y}^b_k)$, where $\mathbf{x}_k = (x_{1k}, \ldots, x_{mk})$, $\mathbf{y}^g_k = (y^g_{1k}, \ldots, y^g_{sk})$, and $\mathbf{y}^b_k = (y^b_{1k}, \ldots, y^b_{qk})$. This vector of inputs and outputs can be viewed as the current or pre-production inputs and output of DMU-k. The inputs and outputs of each DMU-k in S will be changed after producing the additional production quantity. We refer to this as the post-production inputs and outputs of each DMU-k and denote it by $(\mathbf{x}_k + \alpha_k, \mathbf{y}^g_k + \beta_k, \mathbf{y}^b_k + \gamma_k)$, where $\boldsymbol{\alpha}_k = (\alpha_{1k}, \cdots, \alpha_{mk})$, $\boldsymbol{\beta}_k = (\beta_{1k}, \ldots, \beta_{sk})$, and $\boldsymbol{\gamma}_k = (\gamma_{1k}, \ldots, \gamma_{qk})$ are unknowns or decision variables indicating the level of inputs, good outputs and undesirable outputs, respectively, required for producing the additional outputs. Specifically, α_{ik} is the additional level of the ith input that is required for the kth DMU to produce β_{rk} additional (rth) good output and γ_{pk} is the associated generated (pth) undesirable output (for each $i=1, \ldots, m, r=1, \ldots, s, p=1, \ldots, q$ and $k \in S$.

We develop the following inverse DEA model to determine the optimal level of undesirable outputs for DMUs in S generated by the desired production quantity.

$$
\begin{aligned}
&\min\ \gamma = (\gamma_{11}, \ldots, \gamma_{q1}, \ldots, \gamma_{1t}, \ldots, \gamma_{qt}) \\
&\quad s.t. \\
&\qquad \sum_{j\in F} \lambda^k_j x_{ij} + \sum_{l\in S} \hat{\lambda}^k_l (\alpha_{il} + x_{il}) - (\alpha_{ik} + x_{ik}) \le 0 \\
&\qquad\quad \forall k \in S,\ i=1,\ldots,m \\
&\qquad \sum_{j\in F} \lambda^k_j y^g_{rj} + \sum_{l\in S} \hat{\lambda}^k_l (\beta_{rl} + y^g_{rl}) - (1+\hat{\beta}_k) \times (\beta_{rk} + y^g_{rk}) \ge 0 \\
&\qquad\quad \forall k \in S,\ r=1,\ldots,s \\
&\qquad \sum_{j\in F} \lambda^k_j y^b_{pj} + \sum_{l\in S} \hat{\lambda}^k_l (\gamma_{pl} + y^b_{pl}) - (1-\hat{\beta}_k) \times (\gamma_{pk} + y^b_{pk}) = 0 \\
&\qquad\quad \forall k \in S,\ p=1,\ldots,q \\
&\qquad \sum_{j\in F} \lambda^k_j + \sum_{l\in S} \hat{\lambda}^k_l = 1 \qquad \forall k \in S \\
&\qquad \sum_{k\in S} \beta_{rk} = \hat{y}^g_r \quad r=1,\ldots,s \\
&\qquad \alpha_{ik} \ge 0,\ \beta_{rk} \ge 0,\ \gamma_{pk} \ge 0\ \forall k \in S,\ i=1,\ldots,m, \\
&\qquad\quad r=1,\ldots,s,\ p=1,\ldots,q \\
&\qquad \lambda^k_j \ge 0,\ \hat{\lambda}^k_l \ge 0\ \forall k, l \in S,\ \forall j \in F
\end{aligned}
\tag{2}
$$

where $\boldsymbol{\gamma}$ is the vector of undesirable outputs generated by the set of DMUs in S for producing the additional outputs, and t is the number of firms in S. The last constraint in model (2) implies that firms in S should produce $\hat{y}^g_r$ additional good output (for each $r=1$, …, s). In the above model (2), each DMU in S can be represented by a set of efficient DMUs that create the pre-production efficiency frontier, denoted by F, and by a set of DMUs in S. As shown, each DMU-k is represented by a convex combination of some DMUs in F, those with $\lambda^k_j > 0$ ($j \in F$), and some DMUs in S, those with $\hat{\lambda}^k_l > 0$ ($l \in S$). In addition, $(1-\hat{\beta}_k)/(1+\hat{\beta}_k)$ is the efficiency target for DMU-k. The post-production efficiency score of DMU-k will not be deteriorated if and only if $\hat{\beta}_k \le \beta^*_k$, where β^*_k is the optimal value of model (1). Therefore, in the above model (2) we select parameter $\hat{\beta}_k$ such that $\hat{\beta}_k \le \beta^*_k$. In the application of this paper there is only a single undesirable output, or $q=1$, which is GHG emissions

generated by oil and gas firms. In addition, the objective of this paper is to minimize the overall GHG emissions, therefore the above vector optimization inverse DEA model (2) can be simplified to the following single objective model.

$$
\begin{aligned}
&\min \ \gamma_1 + \gamma_2 + \ldots + \gamma_t \\
&s.t. \\
&\sum_{j\in F} \lambda_j^k x_{ij} + \sum_{l\in S} \hat{\lambda}_l^k (\alpha_{il} + x_{il}) - (\alpha_{ik} + x_{ik}) \le 0 \quad \forall k \in S,\ i = 1, \ldots, m \\
&\sum_{j\in F} \lambda_j^k y_{rj}^g + \sum_{l\in S} \hat{\lambda}_l^k (\beta_{rl} + y_{rl}^g) - (1 + \hat{\beta}_k) \times (\beta_{rk} + y_{rk}^g) \ge 0 \\
&\forall k \in S,\ r = 1, \ldots, s \\
&\sum_{j\in F} \lambda_j^k y_j^b + \sum_{l\in S} \hat{\lambda}_l^k (\gamma_l + y_l^b) - (1 - \hat{\beta}_k) \times (\gamma_k + y_k^b) = 0 \quad \forall k \in S \\
&\sum_{j\in F} \lambda_j^k + \sum_{l\in S} \hat{\lambda}_l^k = 1 \qquad \forall k \in S \\
&\sum_{k\in S} \beta_{rk} = \hat{y}_r^g \quad r = 1, \ldots, s \\
&\alpha_{ik} \ge 0,\ \beta_{rk} \ge 0,\ \gamma_k \ge 0\ \forall k \in S,\ i = 1, \ldots, m,\ r = 1, \ldots, s \\
&\lambda_j^k \ge 0,\ \hat{\lambda}_l^k \ge 0\ \forall k, l \in S,\ \forall j \in F
\end{aligned}
\tag{3}
$$

where γ_k is the level of GHG emissions generated by DMU-k for producing the additional output (for each $k \in S$). The objective of the inverse DEA model (3) is to minimize the overall amount of the GHG emissions that will be generated by the selected firms in S for producing the additional desirable outputs $(\hat{y}_1^g, \ldots, \hat{y}_s^g)$.

It is worth noting that the GHG emissions of a firm is a by-product of its production process and hence minimizing some inputs does not necessarily result in minimum the level of GHG emissions. Minimizing some inputs would be a possible objective function after we know the level of GHG emissions that is allowable. In addition, the main concern in this paper is on the environment issue of minimizing GHG emissions and hence the objective function is to minimize the overall GHG emissions.

Theorem 1. Assume there is no frontier movement after producing additional outputs by the selected DUMs in S. Then, the non-linear programming inverse DEA model (3) can be simplified to the following linear programming model.

$$
\begin{aligned}
&\min \ \gamma_1 + \gamma_2 + \ldots + \gamma_t \\
&s.t. \\
&\sum_{j\in F} \lambda_j^k x_{ij} - (\alpha_{ik} + x_{ik}) \le 0 \quad \forall k \in S,\ i = 1, \ldots, m \\
&\sum_{j\in F} \lambda_j^k y_{rj}^g - (1 + \hat{\beta}_k) \times (\beta_{rk} + y_{rk}^g) \ge 0 \quad \forall k \in S,\ r = 1, \ldots, s \\
&\sum_{j\in F} \lambda_j^k y_j^b - (1 - \hat{\beta}_k) \times (\gamma_k + y_k^b) = 0 \quad \forall k \in S \\
&\sum_{j\in F} \lambda_j^k = 1 \qquad \forall k \in S \\
&\sum_{k\in S} \beta_{rk} = \hat{y}_r^g \quad r = 1, \ldots, s \\
&\alpha_{ik} \ge 0,\ \beta_{rk} \ge 0,\ \gamma_k \ge 0\ \forall k \in S,\ i = 1, \ldots, m,\ r = 1, \ldots, s \\
&\lambda_j^k \ge 0\ \forall k, l \in S,\ \forall j \in F
\end{aligned}
\tag{4}
$$

Proof. We show that the assumption of no post-production frontier movement is sufficient to simplify the non-linear model (3) into the linear programming model (4). Assuming no post-production frontier movement implies that the position of DMUs in S after producing the desired outputs will be on or below the pre-production frontier. In either case, DMUs in S can be presented by efficient DMUs defining the pre-production frontier. This fact implies that there is always an optimal solution of the non-linear programming model (3) that $\hat{\lambda}_l^k = 0$ for all $k,\ l \in S$. This completes the proof. □

We discuss a post-production frontier change scenario in Section 5.

Table 1
Number of firms in the final sample by country and year.

	2011	2012	2013	2014	2015
Canadian companies	13	14	12	13	12
US companies	10	10	10	11	11
All companies	23	24	22	24	23

4. Application: data collection

This study utilizes a data set of oil and natural gas producing companies that self-report GHG emissions through the CDP. The CDP is a not-for-profit charity that builds and maintains the world's largest collection of self-reported GHG emissions data. We focus on the energy sector as it accounted for roughly 35% of the global anthropogenic GHG emissions in 2010.[3] Furthermore, while we struggle on a global level to reduce GHG emissions, oil production is expected to grow in Canada by 35% by 2035.[4] Taken together, oil and gas represents a high GHG emitting growth industry that could immensely benefit from improved efficiency.

The initial data set includes all US and Canadian energy sector companies that disclosed scope 1[5] GHG emissions through the CDP between 2011 and 2015. Scope 1 GHG emissions are hand collected from the annual country climate change reports produced by the CDP. The CDP's US edition reports include companies listed on the S&P 500 and the Canadian reports include companies listed on the TSX. The initial data set included 40 North American oil and gas companies with 161 firm-year observation. The 40 companies consisted of 24 Canadian companies and 16 US firms. From the initial sample, 11 companies, with 45 firm-year observation needed to be removed due to a lack of available production information. Of the removed companies, 4 were US oil and gas companies and 7 were Canadian oil and gas companies. The final sample includes 29 North American integrated oil and gas companies with 116 firm-year observations. Table 1 displays the number of oil and gas companies in our dataset for each year. As displayed in Table 1, there are 22–24 firms in the final sample for each year.

Several factors are related to the production of GHG emissions by oil and gas companies. Sueyoshi and Wang (2014) identify drilling activities as being closely related to GHG emissions production. We follow Sueyoshi and Wang (2014) by including the number of new wells drilled each year to our tests. Similar to Sueyoshi and Wang (2014) we also include employees, total assets and capital expenditures. Employees and total assets are proxies for size. Larger firms may have access to more resources to mitigate GHG emissions. Total assets and capital expenditures are further justified as they represent the capital assets that will be used in the production process and new spending that may be invested in GHG mitigation technologies. For US firms, information pertaining to the above mentioned variables and production were hand collected from the 10-K forms retrieved from the Electronic Data Gathering, Analysis, and Retrieval system (EDGAR). For Canadian firms, this information was hand collected from Annual Information Forms and Annual Reports retrieved from the System for Electronic Document Analysis and Retrieval (SEDAR). The following summarizes the data set used in this study. The first two parameters are our outputs, GHG emissions as an undesirable output and

[3] IPCC (2014). Climate Change 2014 Mitigation of Climate Change. <http://www.ipcc.ch/report/ar5/wg3/>. Retrieved: 8/7/2008.

[4] CBC (2018). Canadian total oil production expected to rise by a third by 2035. <https://www.cbc.ca/news/business/oil-production-canada-1.4702433>. Retrieved: 8/7/2018.

[5] Scope 1 emissions are the emissions produced by sources owned by the reporting organization. They are also referred to as direct emissions.

Table 2
Data and efficiency scores of 23 oil and gas firms (Year 2015).

DMUs	Employees	Cap exp	Total assets	Wells	Production	GHG	Efficiency
DMU-1	560	542	4267	61	42	763,457	1.0000
DMU-2	294	376	3948	82	31	1284,090	1.0000
DMU-3	3985	1233	18,553	772	490	5564,499	1.0000
DMU-4	1005	1155	12,672	700	60	2024,889	0.4257
DMU-5	10,765	4465	38,721	3	367	7412,436	1.0000
DMU-6	2891	2232	15,644	400	248	4474,446	0.7366
DMU-7	588	493	2581	47	39	858,929	1.0000
DMU-8	5552	2162	23,779	135	239	11,260,000	1.0000
DMU-9	5700	2586	31,055	46	274	10,711,614	1.0000
DMU-10	449	132	3274	7	26	1020,934	1.0000
DMU-11	12,750	6220	77,527	362	369	18,957,327	0.3978
DMU-12	721	350	3028	36	20	511,213	1.0000
DMU-13	5800	5888	46,414	608	305	11,807,749	0.3300
DMU-14	3860	4748	18,842	467	195	7100,000	0.3229
DMU-15	58,178	29,504	266,103	1866	1955	55,746,124	0.8439
DMU-16	15,900	10,050	97,484	498	580	26,039,254	0.5771
DMU-17	6600	4735	29,532	420	248	5925,440	0.5490
DMU-18	2760	5013	26,975	489	209	6723,280	0.3643
DMU-19	73,500	26,490	336,758	1210	3135	121,000,000	1.0000
DMU-20	2770	4049	34,195	196	136	5561,176	0.4234
DMU-21	1111	1419	4768	139	55	982,304	0.8864
DMU-22	2395	2852	24,196	219	130	2352,253	0.7347
DMU-23	11,100	5272	43,437	671	244	10,400,000	0.2703

production as a good output, and the rest four are the selected inputs.

(a) *GHG emissions:* Total direct/scope 1 GHG emissions as self-reported through the CDP for each year. Direct/scope 1 emissions represent all GHG emissions that are produced directly by the firm. It does not include emissions produced from purchased electricity or any other activity which constitutes the firm's supply chain.
(b) *Production:* This value is taken from the Annual Information Forms provided by Canadian firms or the other information disclosed in 10-K reports by US firms for each year. It is measured in barrels of oil equivalent with 1 barrel being equal to 6000 cubic feet of natural gas.
(c) *Wells:* This value is retrieved from the Annual Information Forms provided by Canadian firms or the other information disclosed in 10-K reports by US firms for each year. The value is the net wells drilled. In other words, it represents the firm's fractional ownership in all wells drilled for each year.
(d) *Employees:* This value is taken from the Annual Reports provided by Canadian firms or the other information disclosed in 10-K reports for US firms for each year. It is a proxy for size. Larger firms may have access to more resources to allow for the mitigation of GHG emissions.
(e) *Capital expenditures:* This value is taken from the firm's statement of cash flows. It is included as it may represent the development of "greener" business processes. All values are translated into US dollars using the December 31 exchange rate for each applicable year.
(f) *Total assets:* This value is collected from the firm's balance sheet. It is included as a proxy for size. It also represents the scope of capital assets that are being employed in the business process. All values are translated into US dollars using the December 31 exchange rate for each applicable year.

5. Result discussion

We use the directional distance DEA model (1) to assess the performance of the oil and gas firms for a period of 5 years (2011–2015). For instance, Table 2 presents the 23 firms each measured based on four inputs and two outputs. The two outputs are production which is the desirable output and GHG emissions as the undesirable output. The table also gives the efficiency score of the DMUs obtained from model (1). Without loss of generality, we omitted the name of the firms.

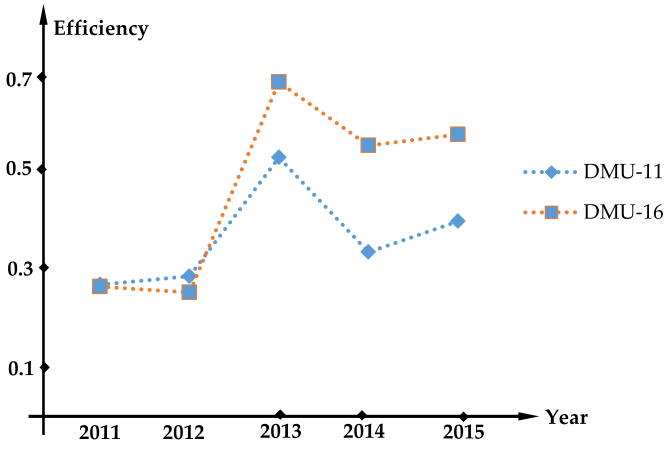

Fig. 1. Efficiency trend for two firms.

Table 3
Optimal GHG emissions.

Production	β_{11}^*	β_{16}^*	γ_{11}^*	γ_{16}^*
$\hat{y}^g = 95$	57.12302	37.87698	0.0000	0.0000
$\hat{y}^g = 110$	57.12302	52.87698	0.0000	205,546.105
$\hat{y}^g = 150$	57.12302	92.87698	0.0000	3230,612.464

To illustrate the applicability of the inverse DEA model, we start with two firms, DMU-11 and DMU-16. As Fig. 1 illustrates, these two firms were inefficient during the years 2011–2015 and consequently dominated by other firms in the market indicating a huge capacity to reduce GHG emissions generated by these two firms.

Let us assume $S = \{\text{DMU} - 11,\ \text{DMU} - 16\}$. The following table presents three different scenarios of implementing the proposed inverse DEA model (4) for the selected DMUs 11 and 16, keeping their efficiency score unchanged.

The first row in Table 3 presents the optimal production quantity by each firm for producing an additional 95 million barrels of

oil equivalents. This production target is equivalent to almost 10% of production of the two firms in 2015. As shown the minimum GHG emissions for each firm, and hence the minimum overall GHG emissions, is zero in this case, indicating there are feasible points in the production possibility set that can produce the desired production quantity with at most the same GHG emissions generated by the two firms in 2015. The next two rows in Table 3 present the optimal solutions of the inverse DEA model (4) associated with two different production quantities. As shown in the last row, for a desired production quantity of $\hat{y}^g = 150$ which is almost 15% of total production, the two firms should produce $\beta_{11}^* = 57.12302$ and $\beta_{16}^* = 92.87698$. In this case, the minimum GHG emissions generated by the two firms is 3230,612.464.

The scenario analysis gives indication that the oil and gas sector has room to increase production without increasing overall GHG emissions even free of a technological shift. An examination of Table 2 reveals that roughly 57% of the firms examined are not on the efficiency frontier. These inefficient DMUs as well as efficient ones all have the potential to increase their production without impacting overall GHG emissions. Alternatively, a space exists for efficient DMUs to acquire inefficient DMUs. By taking advantage of their existing eco-efficiencies, they could increase their production without impacting overall industry GHG emissions.

Now, we illustrate the proposed inverse DEA method by considering a frontier change. The efficiency frontier in DEA is identified by its efficient DMUs. We employ the inverse DEA model (4) for the same set of DMUs, 11 and 16, for a target of an additional production of 150 million barrels of oil incorporating the frontier change. A frontier change may occur by adding efficient firms in 2014 into the observations of 2015. We re-run model (4) by adding the oil and gas firms that were efficient in 2014. In this case, we obtain the following optimal production quantities and GHG emissions.

$$\beta_{11}^* = 58.6,\ \beta_{16}^* = 91.4,\ \gamma_{11}^* = 0,\ \gamma_{16}^* = 0$$

Unlike the optimal solution shown in the third row of Table 3, the minimum GHG emissions is now zero. This simply indicates that the competition can be even tougher for the chosen firms in the case of a frontier change. The two firms were allowed to produce some GHG emissions for an additional production of 150 units when there is no frontier change however, the minimum GHG emissions for the same quantity is zero when we expand the production possibility set.

The proposed inverse DEA approach in this paper can be used to solve such inverse problems: for a set of selected firms what is the maximum achievable production quantity that could be produced by the firms with no additional GHG emissions? The following model solves this inverse problem:

$$\begin{aligned}
&\max\ q\\
&s.t.\\
&\sum_{k\in S}\beta_k = q\\
&\sum_{j\in F}\lambda_j^k x_{ij} + \sum_{l\in S}\hat{\lambda}_l^k(\alpha_{il}+x_{il}) - (\alpha_{ik}+x_{ik}) \le 0 \quad \forall k\in S,\ i=1,\ldots,m\\
&\sum_{j\in F}\lambda_j^k y_j^g + \sum_{l\in S}\hat{\lambda}_l^k(\beta_l + y_l^g) - (1+\hat{\beta}_k)\times(\beta_k + y_k^g) \ge 0 \quad \forall k\in S\\
&\sum_{j\in F}\lambda_j^k y_j^b + \sum_{l\in S}\hat{\lambda}_l^k y_l^b - (1-\hat{\beta}_k)\times y_k^b = 0 \quad \forall k\in S\\
&\sum_{j\in F}\lambda_j^k + \sum_{l\in S}\hat{\lambda}_l^k = 1 \qquad \forall k\in S\\
&\alpha_{ik}\ge 0,\ \beta_k \ge 0\ \forall k\in S,\ i=1,\ldots,m\\
&\lambda_j^k \ge 0,\ \hat{\lambda}_l^k \ge 0\ \forall k,l\in S,\ \forall j\in F
\end{aligned} \tag{5}$$

The inverse DEA model (5) can be of great use to policy makers in terms of helping them determine which firms can improve their eco-efficiency in the current technological environment. More precisely, if the optimal value of model (5) is zero then the set of firms in S cannot produce additional barrels of oil equivalent without increasing their GHG emissions. On the other hand, if the optimal value is positive then at least one of the firms in S can produce some additional barrels of oil equivalent without generating more GHG emissions. It should be this set of firms with the potential to increase their production free of additional GHG emissions that should be the target of GHG emission reduction regulations. The first row in Table 3 indicates that DMU-11 and DMU-16 could produce 95 units of additional production without generating more GHG emissions. They can produce more barrels of oil equivalent without producing more GHG emissions. By extension, they should be capable of producing their current level of oil equivalent with less GHG emissions. To this effect, they are capable of improving performance in response to fines or carbon taxes and thus they should be the focus of regulators. Efficient firms on the other hand should be taxed or fined with caution. These firms should be rewarded for their efforts to minimize their GHG emissions rather than being punished. We suggest that regulations might be more effective at minimizing GHG emissions if they targeted the amount a firm could produce without increasing its GHG emissions as opposed to their actual GHG emissions. This would shift the costs of dealing with climate change away from the firms that are currently doing all that they can to minimize their GHG emissions to those firms that are not. Since the regulations wouldn't impact the cost structure of efficient firms, costs passed to consumers would also likely be lower.

The proposed inverse DEA method in this paper can be used to obtain the minimum sustainable increase of undesirable output associated with the increase of desirable output. This can be derive by looking at the minimum GHG emissions of each firm for certain increase of production over time. For instance, if a firm can increase its production by 5% for a period of time, what would be the sustainable increase of its GHG emissions over that time? The proposed inverse DEA model can be used to estimate the corresponding GHG emissions.

To our knowledge, this paper is the first attempt in the literature to minimize the overall GHG emissions in oil and gas firms through an inverse notion. The only existing study in the literature is the work of Emrouznejad et al. (2018) that solves a different inverse problem for optimal allocation of CO_2 emissions quota in Chinese manufacturing industries.

6. Concluding remarks

As the signs of climate change become more prevalent, the need to lower our anthropogenic greenhouse gas (GHG) emissions becomes more important. Yet it would seem haphazard to race towards GHG emissions reductions without careful consideration of its economic impact. In this study, we seek to find a way to lower GHG emissions in a sector that continues to grow. We develop a novel inverse data envelopment analysis (DEA) model for minimizing the overall GHG emissions for producing a certain level of desirable outputs by a set of chosen firms.

The results of the application of our model to the oil and gas sector reveals some important insights for managers and environmental regulators. We found that roughly 57% of the firms we examined are inefficient. These firms have room to produce more barrels of oil equivalent without increasing their GHG emissions. Managers of these firms have the opportunity to improve their eco-efficiencies to either lower their current GHG emissions levels, or to produce more without additional GHG emissions. Regulators on the other hand, can use our proposed model to identify firms that can improve their eco-efficiency with currently existing technologies. We recommend that regulators should consider targeting

a firm's potential to produce more free from additional GHG emissions as opposed to their actual GHG emissions. This would incentives eco-efficiency improvements without punishing those firms that are already doing all that they can to lower GHG emissions or consumers. While it might not fully address our GHG emissions reduction needs, it represents low bearing fruit that would be a beneficial, and fair, start.

While we apply our proposed model to a subset of firms in the oil and gas industry in both Canada and the US, future studies should extend this application to other industries and should include more countries. It would also be beneficial to expand the sample in the oil and gas industry. If possible, it would be ideal to include all firms. A more complete picture would help estimate how much production could be increased without increasing GHG emissions. Alternatively, a minimum level of GHG emissions could be determined for the global demand of oil and gas. This would create a more complete picture of how close we can come to addressing our GHG emissions reduction needs in our current technological environment without lowering our oil and gas requirements.

Further research is also required to extend the proposed method in this paper to develop a carbon tax strategy based on minimizing GHG emissions. While we suggest that regulators should focus on targeting GHG emission inefficiencies, it would be haphazard to assume that it would result in sufficient GHG emission reductions to fully address climate change. The proposed method should be extended to develop a carbon tax strategy that incentivises both efficient and inefficient firms to lower GHG emissions to the desired levels.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.eswa.2018.12.058.

References

Amin, G. R., Emrouznejad, A., & Gattoufi, S. (2017). Modelling generalized firms' restructuring using inverse DEA. *Journal of Productivity Analysis, 48*(1), 51–61.

Amin, G. R., & Al-Muharrami, S. (2018). A new inverse DEA model for merger with negative data. *IMA Journal of Management Mathematics, 29*(2), 137–149.

Chung, Y. H., Färe, R., & Grosskopf, S. (1997). Productivity and undesirable outputs: A directional distance function approach. *Journal of Environmental Management, 51*(3), 229–240.

Emrouznejad, A., & Yang, G. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences, 61*, 4–8.

Emrouznejad, A., Yang, G., & Amin, G. R. (2018). A novel inverse DEA model with application to allocate the CO_2 emissions quota to different regions in Chinese manufacturing industries. *Journal of the Operational Research Society* In press, https://doi.org/10.1080/01605682.2018.1489344.

Eyni, M., Tohidi, G., & Mehrabeian, S. (2017). Applying inverse DEA and cone constraint to sensitivity analysis of DMUs with undesirable inputs and outputs. *Journal of the Operational Research Society, 68*(1), 34–40.

Gattoufi, S., Amin, G. R., & Emrouznejad, A. (2014). A new inverse DEA method for merging banks. *IMA Journal of Management Mathematics, 25*, 73–87.

Ghiyasi, M. (2017). Inverse DEA based on cost and revenue efficiency. *Computers and Industrial Engineering, 114*, 258–263.

Jahanshahloo, G. R., Soleimani-Damaneh, M., & Ghobadi, S. (2015). Inverse DEA under inter-temporal dependence using multiple-objective programming. *European Journal of Operational Research, 240*(2), 447–456.

Lertworasirikul, S., Charnsethikul, P., & Fang, S. C. (2011). Inverse data envelopment analysis model to preserve relative efficiency values: The case of variable returns to scale. *Computers & Industrial Engineering, 61*(4), 1017–1023.

Lim, D. J. (2016). Inverse DEA with frontier changes for new product target setting. *European Journal of Operational Research, 254*(2), 510–516.

Lin, H. T. (2010). An efficiency-driven approach for setting revenue target. *Decision Support Systems, 49*(3), 311–317.

Molinos-Senante, M., Hernandez-Sancho, F., Mocholi-Arce, M., & Sala-Garrido, R. (2014). Economic and environmental performance of wastewater treatment plants: Potential reductions in greenhouse gases emissions. *Resource and Energy Economics, 38*, 125–140.

Mukherjee, K. (2010). Measuring energy efficiency in the context of an emerging economy: The case of Indian manufacturing. *European Journal of Operational Research, 201*, 933–941.

Riccardi, R., Oggioni, G., & Toninelli, R. (2012). Efficiency analysis of world cement industry in presence of undesirable output: Application of data envelopment analysis and directional distance function. *Energy Policy, 44*, 140–152.

Shetty, U., Pakkala, T. P., & Mallikarjunappa, T. (2012). A modified directional distance formulation of DEA to assess bankruptcy: An application to IT/ITES companies in India. *Expert Systems with Applications, 39*(2), 1988–1997.

Sueyoshi, T., & Goto, G. (2014a). Photovoltaic power stations in Germany and the United States: A comparative study by data envelopment analysis. *Energy Economics, 42*, 271–288.

Sueyoshi, T., & Goto, G. (2014b). DEA radial measurement for environmental assessment: A comparative study between Japanese chemical and pharmaceutical firms. *Applied Energy, 115*, 502–513.

Sueyoshi, T., & Wang, D. (2014). Sustainability development for supply chain management in U.S. petroleum industry by DEA environmental assessment. *Energy Economics, 46*, 360–374.

Toloo, M., Allahyar, M., & Hančlová, J. (2018). A non-radial directional distance method on classifying inputs and outputs in DEA: Application to banking industry. *Expert Systems with Applications, 92*, 495–506.

Wei, Q., Zhang, J., & Zhang, X. (2000). An inverse DEA model for inputs/outputs estimate. *European Journal of Operational Research, 121*(1), 151–163.

Yan, H., Wei, Q., & Hao, G. (2002). DEA models for resource reallocation and production input/output estimation. *European Journal of Operational Research, 136*(1), 19–31.

Vlontzos, G., Niavis, S., & Manos, B. (2014). A DEA approach for estimating the agricultural and environmental efficiency of EU countries. *Renewable and Sustainable Energy Reviews, 40*, 91–96.

Zhou, G., Chung, W., & Zhang, Y. (2014). Measuring energy efficiency performance of China's transport sector: A data envelopment analysis approach. *Expert Systems with Applications, 41*(2), 709–722.